

# IQB Assignment 2

Team name : skrrrrttt.mp3

## Team Members

Anmol Kumar (2018382)

Tejas Dubhir(2018110)

Rakshit Singh(2018079)

For this assignment, we applied the SVM prediction model using the five-fold cross validation method. We carried out this prediction on various input parameters applied on the amino acid sequences which are listed as follows :

- Amino Acid Composition
- Dipeptide Composition
- Tripeptide Composition
- Hydrophobicity

We found out that our best score is **0.92717** using the Tripeptide Composition of amino acid sequences.

## Execution Information

We implemented our code on the Google Colab Platform. We have submitted the .ipynb file for the same. Follow the following steps to execute the code.

- Open [colab.research.google.com](https://colab.research.google.com)
- Open the **final\_tripeptide\_svm.ipynb** which is included in the submission.
- After this, upload the **train.csv** and **test.csv** in the root directory of the notebook. These files are required by the program to read the IDs, sequences and labels.
- Run the program from **Runtime > Run all or using the shortcut Ctrl + F9**
- The final output file named **tripeptide\_svm\_pred.csv** is generated in the root directory of the notebook.